Measures of Association

Crosstabs & Chi-Square - Nominal Level Variables

Measures how different frequencies / percentages are across groups for nominal levels of measurement

Analyze → Descriptive Statistics → Crosstab

1. Place your dependent variable in the Row box
2. Place your independent variables in the Column box
3. Place any "control" variables in the layer box – These may be other variables you want to sort your dependent and independent variables by
4. Click on Statistics
   - Select **Chi-square**
   - For nominal level variables, select **Lambda** and **Phi & Cramer's V**
5. Click Ok to run the output

Output

"Crosstabulation"

- Check the crosstab first to see if there are any obvious relationships

"Chi-Square Tests"

- **Pearson Chi-Square** – check value
  → if the observed data perfectly fit the null hypothesis it would be 0
  → As the number increases, the null hypothesis becomes less likely
- **Asymp. Sig. (2 sided)** - Because this is the p-value it will tell you the probability of obtaining the test statistic by chance –want this to be less than 5% (.05) of the time
  → *if chi-square is not significant, can stop here*

"Directional Measures"

→ *look only at the row that contains your dependent variable for Lambda*

- **Lambda value** – this is a predictive measure so you can make the same interpretations as explained above
  → Lambda is a ***PRE (Proportional Reduction in Error)*** - This is a measurement of strength of a relationship - tells you the extent to which the values of the independent variable predict the values of the dependent variable
    - A value close to 0 indicates that the independent variable provides little predictive help and the relationship is weak
    - Values closer to the poles (-1 and +1) indicate that the independent variable provides a lot of help in predicting the dependent variable. The relationship is strong.
    - If the **Approx. Sig.** value is significant then you can use Lambda as a predictability value. i.e. if the relationship is significant and the Lambda number were .201 – predicts 20.1% of the time.

Note:

* Lambda has a tendency to underestimate the strength of a relationship, especially when one of the variables has low variation (e.g. gender), results in a

.000 value → If chi-square was significant, but is no value for Lambda then scroll down to the next box called Symmetric measures

"Symmetric Measures"

- **Cramer's V –** a 0 value indicates no relationship, 1 indicates a perfect relationship
  → This number is NOT a predictive number because it is not a PRE score, however it shows the relative strength of the relationship between the variables.
  → You would simply report the numbers as decimals – not percentages.

Crosstabs & Chi-Square - Ordinal Level Variables

Analyze → Descriptive Statistics → Crosstab

1. Place your dependent variable in the Row box
2. Place your independent variables in the Column box
3. Place any "control" variables in the "layer" box – These may be other variables you want to sort your dependent and independent variables by
4. Click on Cells → Under Percentages click on Columns and Total (Row percentage optional) → continue
5. Now click on Statistics
   - Select **Chi-square**
   - For ordinal level variables, select Kendall's Tau-b OR Tau-c
     - If the table is square (both variables have the same number of categories) → use **Kendall's Tau-b**
     - If the table is not square (variables have an unequal number of categories) → use **Kendall's Tau-c**
6. Click Ok to run the output

Output

"Crosstabulation"

- Check the crosstab first to see if there are any obvious relationships

"Chi-Square Tests"

- **Pearson Chi-Square** – check value
  → if the observed data perfectly fit the null hypothesis it would be 0
  → As the number increases, the null hypothesis becomes less likely
- **Asymp. Sig. (2 sided)** - Because this is the p-value it will tell you the probability of obtaining the test statistic by chance –want this to be less than 5% (.05) of the time
  → *if chi-square is not significant, can stop here* "Symmetric Measures"
- The **Kendall's Tau-b or –c** number under the value column reports the strength of the relationship.
  → Both provide a ***PRE (Proportional Reduction in Error)*** - This is a measurement of strength of a relationship - tells you the extent to which the values of the independent variable predict the values of the dependent variable

- A value close to 0 indicates that the independent variable provides little predictive help and the relationship is weak
- Values closer to the poles (-1 and +1) indicate that the independent variable provides a lot of help in predicting the dependent variable. The relationship is strong.
- If the **Asymp. Sig. (2-sided)** value is significant then you can use the Kendall's Tau number as a predictability value. i.e. if the relationship is significant and the Kendall's number were .201 – you could say that given the variables tested you could predict the outcome with 95% confidence about 20.1% of the time.

→ Both are ***directional measures***: a positive value indicates that increasing values on the independent variable are associated with increasing values on the dependent variable and vice versa

Note:

* Take note of the sign of your Kendall's Tau-b or -c number – SPSS does not take into consideration how your hypothesis is worded so it may give the opposite sign to this value. For example it may show a negative sign for this value when in reality it is a positive relationship. Be sure to look closely at your Crosstab output categories to see what the true direction of the relationship is